Supporting Information

Putnam et al. Bioinformatic identification of genes suppressing genome instability

SI Text

SI Materials and Methods. ***Calculation of p_{nhit}.*** The p_{nhit} *p*-value for mutations observed in *n* of the *N* DNA damaging screens was calculated using the following equation:

$$p_{nhit} = \sum_{i=n}^{N} \text{prob}(i) \qquad \textbf{[S1]}$$

The probability of observing a mutation in *n* screens, prob(*n*), was based on computer simulation combining 1,000,000 trials in which mutations were selected at random. For the simulation of each screen, m_i random mutations were chosen from 4,756 possible mutations (corresponding to the number of strains in the haploid yeast deletion collection), where m_i was the number of mutations observed in the *i*th screen. After randomly simulating all *N* screens, the number mutations selected from 0 to *N* times in each trial was recorded, and these results were summed for all 1,000,000 trials of the *N* screens. The prob(*n*) value was calculated by counting the total number of observations of mutations observed *n* times divided by the total number of observations of mutations hit any number of times.

These random simulations give remarkably consistent probabilities (and hence p-values) that incorporate both the number of screens and the number of genes in each screen regardless of how the calculation is performed. For example, the probability of a mutation being identified in 4 of 155 screens is 0.1354. If we divide the screens into 22 mechlorethamine screens (with 27, 32, 34, 97, 103, 113, 119, 124, 125, 147, 153, 164, 173, 174, 174, 175, 193, 198, 204, 288, 367, and 418 genes in the different mechlorethamine screens) and 133 non-mechlorethamine screens (with variable amounts of genes identified). The probability of identifying a mutation 4 times in the 155 screen by calculating the probabilities of identifying a mutation in *n1* mechlorethamine screens and *n2* non-mechlorethamine screens, where $n1 + n2 = 4$. The probability for each case of *n1* and *n2* is:

$$p_{combined}(n1,n2) = p_{in}(n1)\, p_{out}(n2)$$

where $p_{in}(n1)$ is the probability of finding a gene *n1* times in a mechlorethamine screen and $p_{out}(n2)$ is the probability of finding a gene *n2* times in non-mechlorethamine screens. The sum of all of the independent possibilities (generated by 10 different simulations), 0.135396**,** comes to the same result as the simulation of all 155 different screens, 0.1354 (Table S6). A similar calculation can be performed for any other separation of 155 screens into two sets of screens, indicating that the individual probabilities calculated are consistent and robust.

Determination if a mutation was specific to a particular group of screens. To determine if the distribution of a mutation was caused by specificity to a particular group of screens, *G*, such as those belonging to a particular DNA damaging agent or being performed in a particular laboratory, we used the equation S1 and the simulations as described above to calculate *p*-values. We calculated both p_{in}, the p_{nhit} for the mutation in *G*, and p_{out}, the p_{nhit} for the mutation in *G´*, the group of screens not in *G*. We used these *p*-values to calculate likelihoods that the distribution of a mutation within *G* and *G´* corresponded to noise, specificity to *G*, or significant (common) to both *G* and *G´*.

$$L_{noise} = [\, p_{in} \,]\, [\, p_{out} \,]$$
$$L_{specific} = [\, 1 - p_{in} \,]\, [\, p_{out} \,]$$
$$L_{common} = [\, 1 - p_{in} \,]\, [\, 1 - p_{out} \,]$$

We then calculated the ratios $L_{specific}/L_{noise}$ and $L_{specific}/L_{common}$:

$$L_{specific}/L_{noise} = [\, 1 - p_{in} \,] / [\, p_{in} \,]$$
$$L_{specific}/L_{common} = [\, p_{out} \,] / [\, 1 - p_{out} \,]$$

If both ratios were greater than 1, we flagged the mutation as specific to the group of genes under investigation. For example, *pef1*Δ was observed in 5 of 6 bleomycin screens (p_{in}=0.00) and 2 of 149 (p_{out}=0.71) non-bleomycin screens and was flagged as bleomycin-specific.

Calculation of genetic distance via the composite angle distance. Growth based genetic interactions were combined to form a binary interaction matrix; pairs of genes were scored as "having" or "not having" a growth based interaction with each other when mutated. The resulting binary interaction matrix was used to calculate a genetic distance between each pair of genes. For each pair of genes *A* and *B*, we define M_{01} as the number of genes that only interact with *A*, M_{10} as the number of genes that only interact with *B*, and M_{11} as the number of genes that interact with both. Using these counts, we define a two-dimensional vector $\boldsymbol{v}_{A,B} = (M_{11}, M_{01} + M_{10})$ (Fig. S2A). The angle between $\boldsymbol{v}_{A,B}$ and the *x*-axis, which ranges between 0 and π/2 radians (0 and 90 degrees), was calculated and scaled to generate the genetic distance, which ranges between 0.0 (all interactions shared) and 1.0 (no interactions shared). Thus, the composite angle distance between genes *A* and *B*, CAD(*A,B*), for binary interactions can be defined as:

$$\text{CAD}(A,B) = 2\, (\, atan(\, (M_{01} + M_{10}) / M_{11}) \,) \,/\, \pi \qquad \textbf{[S2]}$$

We note that when individual interactions can be weighted by the strength of the interaction, these weights can be directly included into the construction of the vector $\boldsymbol{v}_{A,B}$.

For pairwise comparisons with binary interaction data, CAD is similar to the commonly used Jaccard distance (57; Fig. S2B,C), which is:

$$\begin{aligned} \text{JD}(A,B) &= (\, |A \cup B| - |A \cap B| \,) / (\, |A \cup B| \,) \\ &= (M_{01} + M_{10}) / (M_{01} + M_{10} + M_{11}) \end{aligned} \qquad \textbf{[S3]}$$

CAD also gives similar results to the cosine distance (Fig. S2B,D), which is defined as follows where $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ are the N-dimensional vectors of interactions for genes *A* and *B*):

$$\begin{aligned} \text{CD}(A,B) &= 1 - \boldsymbol{v}_A \bullet \boldsymbol{v}_B \,/\, (\, |\boldsymbol{v}_A|\, |\boldsymbol{v}_B| \,) \\ &= 1 - M_{11} / (\, sqrt(M_{01} + M_{11})\, sqrt(M_{10} + M_{11}) \,) \end{aligned} \qquad \textbf{[S4]}$$

Importantly, all of these distance measures only take into account reported interactions, which is crucial for analysis of data present in databases like Biogrid, where a measured lack of interaction is not reported. For measuring interactions between two genes, the CAD is equivalent to these measures; however, the formulation of CAD naturally extends to measuring genetic distances between groups of genes so that individual interactions are appropriately weighted.

Calculation of genetic distances between groups of genes. Determining the composite angle distance between the group of genes X, containing of N_X genes, and the group of genes Y, containing N_Y genes, was performed by summing all pairwise vectors between the genes in group X and the genes in group Y to generate vector $\boldsymbol{v}_{TOTAL}$:

$$\boldsymbol{v}_{TOTAL} = \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} \boldsymbol{v}_{i,j} \qquad \textbf{[S5]}$$

The CAD(X,Y) distance is therefore the angle between $\boldsymbol{v}_{TOTAL}$ and the x-axis scaled to be between 0 and 1, and was equivalent to calculating the distance using equation S2, where M_{01}, M_{10}, and M_{11} were taken to be the sum of all individual pairwise comparisons of genes in group X with those in group Y (Fig. S2E).

The advantage of the composite angle distance for handling groups of genes is that the method appropriately scales the effects of individual genes weighted by the number of interactions. Methods that calculate distances by determining, for example, the average, maximum, or minimum of all pairwise interactions, ignore the number of interactions that define these distances. For example, if gene A shares 1 of the 100 interactions of gene $B1$ and 1 of the 1 interactions of genes $B2$ and $B3$, then the pairwise Jaccard distances would be JD($A,B1$)=0.99, JD($A,B2$)=0.00, and JD($A,B3$)=0.00, with an average of 0.33. By the CAD method, $\boldsymbol{v}_{TOTAL}$ = (3, 99), so the distance is 0.98. The CAD score therefore scales distance by interactions and not by genes, which is critical for appropriate weighting when scoring genetic congruence as described below.

Scoring genetic congruence. We scored genetic congruence of each gene in the database against the list of genes of interest using the composite angle distance method. Over 100,000 random simulations were performed to calculate p-values. In order to appropriately calculate p-values, these random simulations had to appropriately account for the effects of genes containing many genetic interactions as well as those containing few genetic interactions. Thus, in each simulation evaluating N real genes, N theoretical genes were generated. Each theoretical gene was constructed to contain the same number of genetic interactions with targets as the corresponding real gene, but the theoretical gene had randomly selected targets. Target selection was weighted by number of genetic interactions of each target. Weighting targets by their number of genetic interactions and constructing theoretical genes with the same number of interactions as the genes of interest appropriately accounted for differences between "hub" genes and "spoke" genes during the simulations.

Clustering. Genes were clustered on the basis of their genetic congruence by the composing angle distance method using agglomerative hierarchical clustering (37) with the modification that new clusters are built at each step from all elements whose best congruence scores are with each other rather than just building a single cluster at each step from the two elements with the best congruence score.

1. Chen C & Kolodner RD (1999) Gross chromosomal rearrangements in *Saccharomyces cerevisiae* replication and recombination defective mutants. *Nat Genet* 23(1):81-85.
2. Huang ME, Rio AG, Nicolas A, & Kolodner RD (2003) A genomewide screen in *Saccharomyces cerevisiae* for genes that suppress the accumulation of mutations. *Proc Natl Acad Sci U S A* 100(20):11529-11534.
3. Hwang JY, Smith S, & Myung K (2005) The Rad1-Rad10 complex promotes the production of gross chromosomal rearrangements from spontaneous DNA damage in *Saccharomyces cerevisiae*. *Genetics* 169(4):1927-1937.
4. Myung K, Datta A, Chen C, & Kolodner RD (2001) *SGS1*, the *Saccharomyces cerevisiae* homologue of *BLM* and *WRN*, suppresses genome instability and homeologous recombination. *Nat Genet* 27(1):113-116.
5. Myung K, Pennaneach V, Kats ES, & Kolodner RD (2003) *Saccharomyces cerevisiae* chromatin-assembly factors that act during DNA replication function in the maintenance of genome stability. *Proc Natl Acad Sci U S A* 100(11):6640-6645.
6. Pennaneach V & Kolodner RD (2004) Recombination and the Tel1 and Mec1 checkpoints differentially effect genome rearrangements driven by telomere dysfunction in yeast. *Nat Genet* 36(6):612-617.
7. Huang ME & Kolodner RD (2005) A biological network in *Saccharomyces cerevisiae* prevents the deleterious effects of endogenous oxidative DNA damage. *Mol Cell* 17(5):709-720.
8. Myung K, Chen C, & Kolodner RD (2001) Multiple pathways cooperate in the suppression of genome instability in *Saccharomyces cerevisiae*. *Nature* 411(6841):1073-1076.
9. Myung K, Datta A, & Kolodner RD (2001) Suppression of spontaneous chromosomal rearrangements by S phase checkpoint functions in *Saccharomyces cerevisiae*. *Cell* 104(3):397-408.
10. Myung K & Kolodner RD (2002) Suppression of genome instability by redundant S-phase checkpoint pathways in *Saccharomyces cerevisiae. Proc Natl Acad Sci U S A* 99(7):4500-4507.
11. Schmidt KH & Kolodner RD (2006) Suppression of spontaneous genome rearrangements in yeast DNA helicase mutants. *Proc Natl Acad Sci U S A* 103(48):18196-18201.
12. Schmidt KH, Wu J, & Kolodner RD (2006) Control of translocations between highly diverged genes by Sgs1, the *Saccharomyces cerevisiae* homolog of the Bloom's syndrome protein. *Mol Cell Biol* 26(14):5406-5420.
13. Smith S, Gupta A, Kolodner RD, & Myung K (2005) Suppression of gross chromosomal rearrangements by the multiple functions of the Mre11-Rad50-Xrs2 complex in *Saccharomyces cerevisiae*. *DNA Repair (Amst)* 4(5):606-617.
14. Banerjee S, Smith S, & Myung K (2006) Suppression of gross chromosomal rearrangements by yKu70-yKu80 heterodimer through DNA damage checkpoints. *Proc Natl Acad Sci U S A* 103(6):1816-1821.
15. Budd ME, Reis CC, Smith S, Myung K, & Campbell JL (2006) Evidence suggesting that Pif1 helicase functions in DNA replication with the Dna2 helicase/nuclease and DNA polymerase delta. *Mol Cell Biol* 26(7):2490-2500.
16. De Piccoli G, et al. (2006) Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. *Nat Cell Biol* 8(9):1032-1034.
17. Motegi A, Kuntz K, Majeed A, Smith S, & Myung K (2006) Regulation of gross chromosomal rearrangements by ubiquitin and SUMO ligases in *Saccharomyces cerevisiae*. *Mol Cell Biol* 26(4):1424-1433.
18. Lee W, et al. (2005) Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet* 1(2):e24.
19. Hillenmeyer ME, et al. (2008) The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* 320(5874):362-365.
20. Aouida M, Page N, Leduc A, Peter M, & Ramotar D (2004) A genome-wide screen in *Saccharomyces cerevisiae* reveals altered transport as a mechanism of resistance to the anticancer drug bleomycin. *Cancer Res* 64(3):1102-1109.
21. Brown JA, et al. (2006) Global analysis of gene function in yeast by quantitative phenotypic profiling. Mol Syst Biol 2:2006 0001.
22. Lum PY, et al. (2004) Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* 116(1):121-137.
23. Parsons AB, et al. (2006) Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* 126(3):611-625.
24. Kitagawa T, Hoshida H, & Akada R (2007) Genome-wide analysis of cellular response to bacterial genotoxin CdtB in yeast. *Infect Immun* 75(3):1393-1402.
25. Giaever G, et al. (2004) Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci U S A* 101(3):793-798.
26. Wu HI, Brown JA, Dorie MJ, Lazzeroni L, & Brown JM (2004) Genome-wide identification of genes conferring resistance to the anticancer agents cisplatin, oxaliplatin, and mitomycin C. *Cancer Res* 64(11):3940-3948.
27. Baldwin EL, Berger AC, Corbett AH, & Osheroff N (2005) Mms22p protects *Saccharomyces cerevisiae* from DNA damage induced by topoisomerase II.

Nucleic Acids Res 33(3):1021-1030.
28. Hartman JLT & Tippery NP (2004) Systematic quantification of gene interactions by phenotypic array analysis. *Genome Biol* 5(7):R49.
29. Bennett CB, et al. (2001) Genes required for ionizing radiation resistance in yeast. *Nat Genet* 29(4):426-434.
30. Game JC, et al. (2003) Use of a genome-wide approach to identify new genes that control resistance of *Saccharomyces cerevisiae* to ionizing radiation. *Radiat Res* 160(1):14-24.
31. Chang M, Bellaoui M, Boone C, & Brown GW (2002) A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. *Proc Natl Acad Sci U S A* 99(26):16934-16939.
32. Hanway D, et al. (2002) Previously uncharacterized genes in the UV- and MMS-induced DNA damage response in yeast. *Proc Natl Acad Sci U S A* 99(16):10605-10610.
33. Begley TJ, Rosenbach AS, Ideker T, & Samson LD (2004) Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping. *Mol Cell* 16(1):117-125.
34. Birrell GW, Giaever G, Chu AM, Davis RW, & Brown JM (2001) A genome-wide screen in *Saccharomyces cerevisiae* for genes affecting UV radiation sensitivity. *Proc Natl Acad Sci U S A* 98(22):12608-12613.
35. Smith S, et al. (2004) Mutator genes for suppression of gross chromosomal rearrangements identified by a genome-wide screening in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 101(24):9039-9044.
36. Kanellis P, et al. (2007) A screen for suppressors of gross chromosomal rearrangements identifies a conserved role for PLP in preventing DNA lesions. *PLoS Genet* 3(8):e134.
37. Xu R & Wunsch II D (2005) Survey of clustering algorithms. *IEEE Transactions on Neural Networks* 16:645-678.
38. Alvaro D, Lisby M, & Rothstein R (2007) Genome-wide analysis of Rad52 foci reveals diverse mechanisms impacting recombination. *PLoS Genet* 3(12):e228.
39. Askree SH, et al. (2004) A genome-wide screen for *Saccharomyces cerevisiae* deletion mutants that affect telomere length. *Proc Natl Acad Sci U S A* 101(23):8658-8663.
40. Gatbonton T, et al. (2006) Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. *PLoS Genet* 2(3):e35.
41. Griffith JL, et al. (2003) Functional genomics reveals relationships between the retrovirus-like Ty1 element and its host *Saccharomyces cerevisiae*. *Genetics* 164(3):867-879.
42. Scholes DT, Banerjee M, Bowen B, & Curcio MJ (2001) Multiple regulators of Ty1 transposition in *Saccharomyces cerevisiae* have conserved roles in genome maintenance. *Genetics* 159(4):1449-1465.
43. Irwin B, et al. (2005) Retroviruses and yeast retrotransposons use overlapping sets of host genes. *Genome Res* 15(5):641-654.
44. Ouspenski, II, Elledge SJ, & Brinkley BR (1999) New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability. *Nucleic Acids Res* 27(15):3001-3008.
45. Yuen KW, et al. (2007) Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci U S A* 104(10):3925-3930.
46. Andersen MP, Nelson ZW, Hetrick ED, & Gottschling DE (2008) A genetic screen for increased loss of heterozygosity in *Saccharomyces cerevisiae*. *Genetics* 179(3):1179-1195.

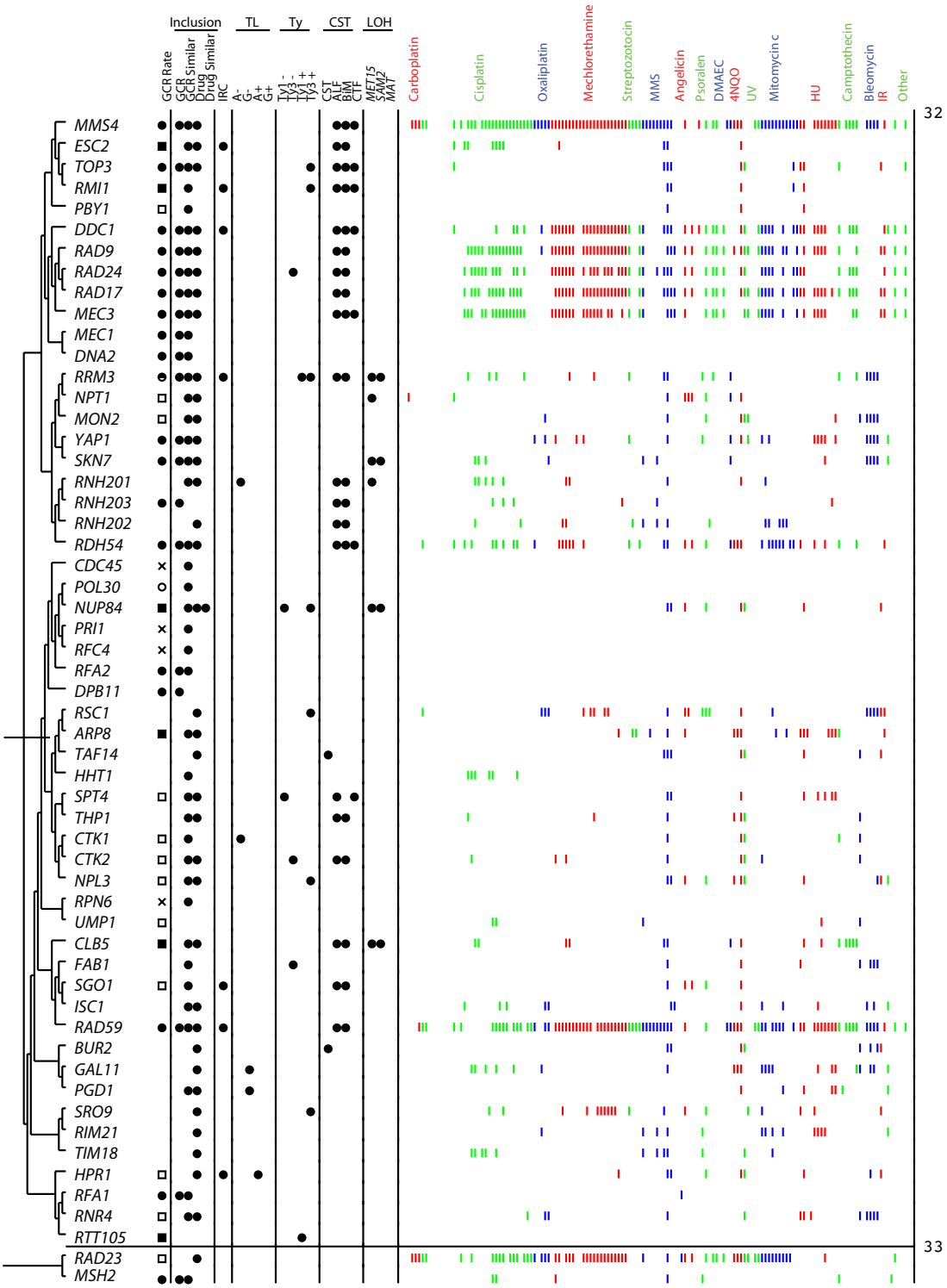

Fig. S1. Annotated view of clusters 32 and 33. The GCR Rate column identifies mutations tested in the GCR assay: circles were previously tested; squares were tested in this study; crosses were essential genes; filled-in symbols increased GCR rates as single mutants; half filled-in symbols only synergistically increased GCR rates in combination with other mutants; and open symbols did not increase GCR rates. "Inclusion" indicates if a gene was identified in the GCR rate (GCR), genetic congruence to GCR genes (GCR similar), DNA damaging agent (Drug), or genetic congruence to DNA damaging agent genes (Drug Similar) stages of the bioinformatics analysis. "IRC" indicates those genes causing Increased Recombination Centers (38). "TL" indicates mutations identified in two telomere-length screens by Askree *et al.* and Gatbonton *et al.* (39, 40) with decreased (A-, G-) or increased (A+, G+) telomere lengths. "Ty" indicates mutations causing decreased (Ty1-, Ty3-) or increased (Ty1+, Ty3+) transposition (41-43). "CST" indicates mutations identified as affecting chromosome stability by several assays (44, 45). LOH indicates mutations increasing loss-of-heterozygosity by several assays (46). Sensitivity to each DNA damaging agents is indicated by vertical bars, with different treatments having alternate colors.

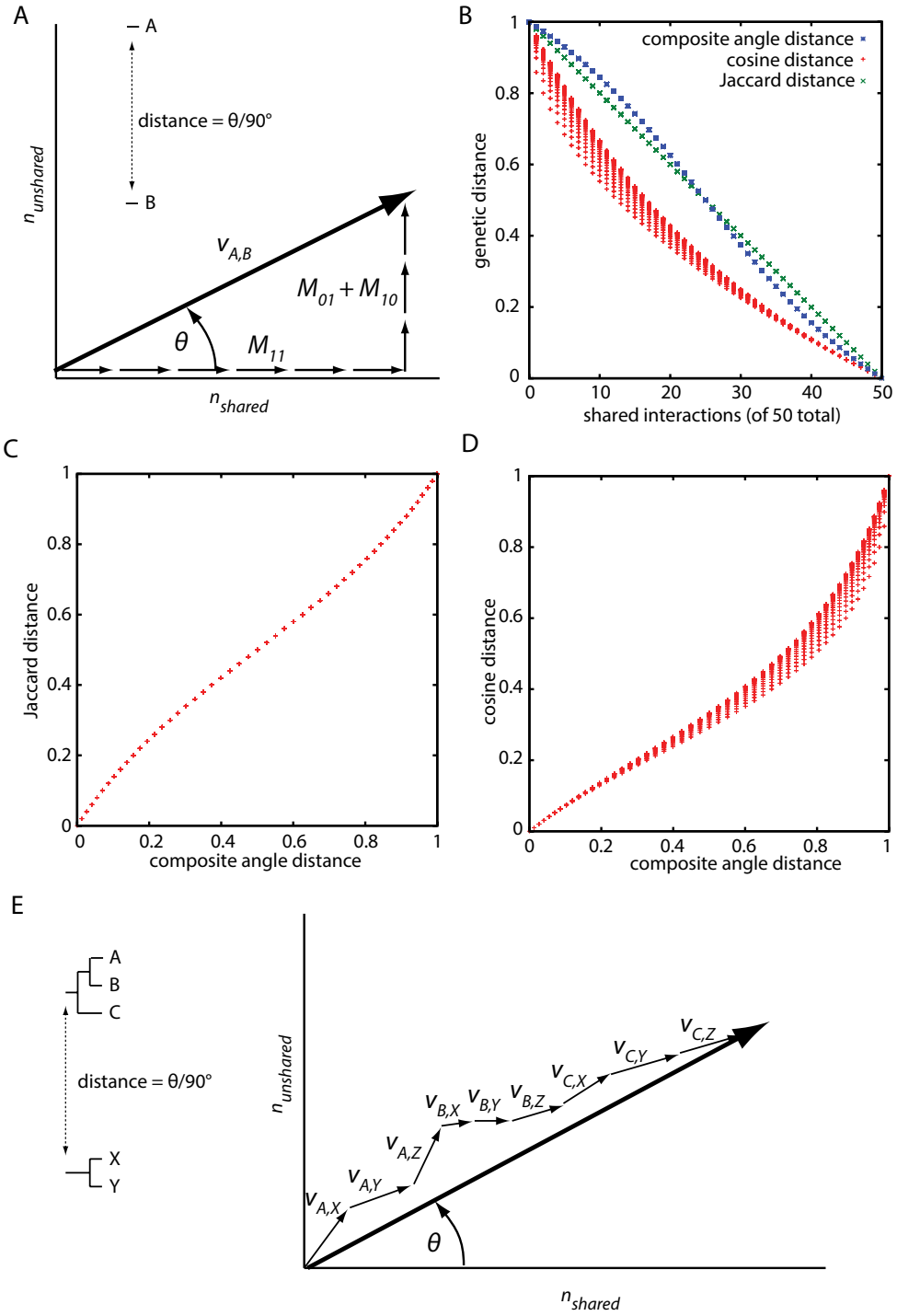

Fig. S2. The composite angle distance measure of genetic similarity. (A) To measure the distance between genes *A* and *B*, we count the number of shared interactions, M_{11}, the number of interactions specific to gene *A*, M_{01}, and the number of interactions specific to gene *B*, M_{10} and construct the vector $\boldsymbol{v}_{A,B} = (M_{11}, M_{01}+M_{10})$. The composite angle distance is defined to be the angle between $\boldsymbol{v}_{A,B}$ divided by 90 degrees. (B) Comparison of the distance measured by the composite angle distance (blue), the cosine distance (red), and the Jaccard distance (green) for all of the cases involving interactions with 50 different targets ($M_{01}+M_{10}+M_{11}=50$). Note that the cosine distance measure is sensitive to the distribution of unmatched interactions, M_{01} and M_{10}, whereas the Jaccard and composite distance measure are not. (C) Plot of the Jaccard distance against the composite angle distance reveals that these measures are very similar. (D) Plot of the cosine distance against the composite angle distance reveals that these measures are also very similar. (E) Extension of the composite angle distance to measure genetic distances between two groups of genes, *A-B-C* and *X-Y*, by summation of individual pair-wise vectors.

Table S1. Mutations that cause increased genomic instability as single mutations and/or cause synergistic increases in genomic instability in combination with other mutations.

Mutation	Systematic name	GCR rate (fold)*	Synergistic GCRs†	Number of DNA Damaging Screens (out of 155)
rad27	*ykl113c*	3.64×10^{-7} (1040)	No	13
pif1	*yml061c*	3.53×10^{-7} (1010)	No	4
rfa1-t33	*yar007c*	3.37×10^{-7} (962)	Yes	1‡
rad50	*ynl250w*	2.30×10^{-7} (657)	Yes	53
mre11	*ymr224c*	1.93×10^{-7} (550)	Yes	27
xrs2	*ydr369c*	1.90×10^{-7} (543)	Yes	26
sic1	*ylr079w*	1.80×10^{-7} (514)	No	3
cac1 (rlf2)	*ypr018w*	1.20×10^{-7} (343)	Yes	6
rad18	*ycr066w*	9.50×10^{-8} (271)	No	74
dpb11-1	*yjl090c*	6.75×10^{-8} (193)	Yes	0‡
pds1	*ydr113c*	6.70×10^{-8} (191)	Yes	0
mus81	*ydr386w*	6.50×10^{-8} (186)	Yes	106
rfc5-1	*ybr087w*	6.05×10^{-8} (173)	Yes	0‡
mms4	*ybr098w*	5.90×10^{-8} (169)	Yes	119
ddc2 (lcd1) sml1	*ydr499w*	5.70×10^{-8} (163)	Yes	0‡
mec1 sml1	*ybr136w*	5.42×10^{-8} (155)	Yes	0‡
dun1	*ydl101c*	4.81×10^{-8} (137)	Yes	43
alo1	*yml086c*	4.70×10^{-8} (134)	No	0
rad5	*ylr032w*	4.40×10^{-8} (126)	Yes	118
rad52	*yml032c*	4.01×10^{-8} (115)	Yes	30
dia2	*yor080w*	3.69×10^{-8} (105)	No	7
cac2	*yml102w*	3.00×10^{-8} (85.7)	No	20
rfa3-n70	*yjl173c*	2.80×10^{-8} (80.0)	No	0‡
mms21-11	*yel019c*	2.80×10^{-8} (80.0)	No	1‡
smc6-9	*ylr383w*	2.70×10^{-8} (77.1)	No	0‡
ufo1	*yml088w*	2.60×10^{-8} (74.3)	No	5
asf1	*yjl115w*	2.50×10^{-8} (71.4)	Yes	12
nse3-2	*ydr288w*	1.90×10^{-8} (54.3)	No	0‡
rfa2-c100	*ynl312w*	1.90×10^{-8} (54.3)	No	0‡
elg1	*yor144c*	1.58×10^{-8} (45.3)	Yes	26
mec3	*ylr288c*	1.39×10^{-8} (39.7)	Yes	66
tsa1	*yml028w*	1.24×10^{-8} (35.4)	Yes	10
sgs1	*ymr190c*	1.20×10^{-8} (34.4)	Yes	41
rad57	*ydr004w*	1.16×10^{-8} (33.3)	Yes	70
chk1	*ybr274w*	1.11×10^{-8} (31.8)	Yes	0
cac3 (msi1)	*ybr195c*	1.10×10^{-8} (31.4)	Yes	35
ogg1	*yml060w*	1.05×10^{-8} (29.9)	Yes	0
top3	*ylr234w*	9.50×10^{-9} (27.1)	No	16
rad53 sml1	*ypl153c*	8.24×10^{-9} (23.5)	Yes	0‡
rtt101	*yjl047c*	7.88×10^{-9} (22.5)	No	57
rad59	*ydl059c*	7.50×10^{-9} (21.4)	Yes	96
dna2-2	*yhr164c*	7.00×10^{-9} (20.0)	No	0‡
rad17	*yor368w*	5.64×10^{-9} (16.1)	Yes	72
rad9	*ydr217c*	5.33×10^{-9} (15.2)	Yes	76
stn1-13	*ydr082w*	5.30×10^{-9} (15.1)	No	0‡
rif2	*ylr453c*	5.00×10^{-9} (14.3)	No	5
cdc50	*ycr094w*	4.80×10^{-9} (13.7)	Yes	45
ydl162c	*ydl162c*	4.59×10^{-9} (13.1)	No	16
rad24	*yer173w*	4.00×10^{-9} (11.4)	Yes	63
shu2	*ydr078c*	3.55×10^{-9} (10.1)	No	46
rad51	*yer095w*	3.50×10^{-9} (10.0)	Yes	63
rnh203	*ylr154c*	2.96×10^{-9} (8.46)	No	6
shu1	*yhl006c*	2.95×10^{-9} (8.43)	No	46
skn7	*yhr206w*	2.91×10^{-9} (8.31)	No	13
exo1	*yor033c*	2.70×10^{-9} (7.71)	No	18
msh2	*yol090w*	2.53×10^{-9} (7.21)	No	7
rad55	*ydr076w*	2.40×10^{-9} (6.86)	No	83
esc1	*ymr219w*	2.30×10^{-9} (6.57)	No	0
yap1	*yml007w*	2.20×10^{-9} (6.29)	No	25
rad54	*ygl163c*	2.07×10^{-9} (5.90)	Yes	59
ddc1	*ypl194w*	2.00×10^{-9} (5.71)	Yes	65
cln2-1	*ypl256c*	1.99×10^{-9} (5.69)	No	2

Mutation	Systematic name	GCR rate (fold)*	Synergistic GCRs†	Number of DNA Damaging Screens (out of 155)
csm2	*yil132c*	1.78×10^{-9} (5.07)	Yes	73
rdh54	*ybr073w*	1.75×10^{-9} (5.00)	Yes	42
lig4 (dnl4)	*yor005c*	1.60×10^{-9} (4.57)	Yes	1
rrm3	*yhr031c*	1.40×10^{-9} (4.00)	Yes	20
siz1	*ydr409w*	1.30×10^{-9} (3.71)	Yes	45
tel1	*ybl088c*	6.71×10^{-10} (1.92)	Yes	9
pol32	*yjr043c*	4.00×10^{-10} (1.14)	Yes	72
tlc1	*tlc1*	3.20×10^{-10} (0.91)	Yes	0
srs2	*yjl092w*	3.15×10^{-10} (0.90)	Yes	88
est2	*ylr318w*	2.17×10^{-10} (0.62)	Yes	10
est1	*ylr233c*	1.50×10^{-10} (0.43)	Yes	5
est3	*yil009c-a*	1.50×10^{-10} (0.43)	Yes	12
siz2 (nfi1)	*yor156c*	1.50×10^{-10} (0.43)	Yes	7

*Rate of the single mutant only, derived by analysis of published rates (1-17). Fold increase over the wild-type rate, $3.5\text{x}10^{-10}$ (1), in parentheses.
†Indicates if synergistic interactions in the GCR assay are known.
‡Single deletion mutations are lethal.

Table S2. Mutations not causing increased GCRs in single-copy sequences.

Mutation name	Systematic name	GCR rate*	Suppress GCR rates†
apn1	*ykl114c*	4.8×10^{-10}	No
apn2	*ybl019w*	3.8×10^{-10}	Yes
bre1	*ydl074c*	9.9×10^{-10}	Yes
bub1	*ygr188c*	4.7×10^{-10}	Yes
bub2	*ymr055c*	3.4×10^{-10}	Yes
bub3	*yor026w*	3.9×10^{-10}	Yes
cdc13-2	*ydl220c*	4.6×10^{-10}	Yes
ctf18	*ymr078c*	3.2×10^{-10}	Yes
ctf8	*yhr191c*	3.5×10^{-10}	Yes
dcc1	*ycl016c*	4.1×10^{-10}	Yes
dna2-1	*yhr164c*	3.5×10^{-10}	No
lif1	*ygl090w*	4.0×10^{-10}	No
lys7	*ymr038c*	5.0×10^{-10}	No
mad2	*yjl030w*	5.6×10^{-10}	Yes
mad3	*yjl013c*	2.4×10^{-10}	Yes
mag1	*yer142c*	4.3×10^{-10}	No
mms2	*ygl087c*	2.6×10^{-10}	No
msh6	*ydr097c*	1.6×10^{-9}	No
ntg1	*yal015c*	6.3×10^{-10}	No
ntg2	*yol043c*	3.0×10^{-10}	No
pol30-119	*ybr088c*	1.0×10^{-9}	Yes
psy3	*ylr376c*	5.0×10^{-10}	No
rad1	*ypl022w*	1.9×10^{-10}	Yes
rad6	*ygl058w*	6.1×10^{-10}	Yes
rad10	*yml095c*	1.0×10^{-10}	Yes
rad30	*ydr419w*	4.3×10^{-10}	No
rev1	*yor346w*	4.6×10^{-10}	Yes
rev3	*ypl167c*	4.1×10^{-10}	Yes
rif1	*ybr275c*	9.9×10^{-10}	No
sir1	*ykl101w*	8.9×10^{-10}	No
sir2	*ydl042c*	2.5×10^{-10}	Yes
sir3	*ylr442c*	5.0×10^{-10}	Yes
sir4	*ydr227w*	8.4×10^{-10}	No
siz1	*ydr409w*	1.3×10^{-9}	Yes
sml1	*yml058w*	3.1×10^{-10}	No
sod1	*yjr104c*	8.8×10^{-10}	No
ubc13	*ydr092w*	1.3×10^{-9}	No
ung1	*yml021c*	3.5×10^{-10}	No
yku70	*ymr284w*	5.4×10^{-10}	Yes
yku80	*ymr106c*	7.8×10^{-10}	Yes

*The wild-type rate is 3.5×10^{-10} (1).

†Indicates if mutation is known to suppress the increased GCR rate of other mutations.

Table S3. Included genome-wide DNA damaging agent screens.

Damaging Agent	Treatment*	Title	Number of Mutations†	Reference
Angelicin	62.5 uM irradiated; HOMD; CT; CLC; MARO	04_03_17_12	165	(18, 19)
	62.5 uM irradiated; HOMD; CT; CLC; MARO	04_03_25_06	64	(18, 19)
	62.5 uM irradiated; HOMD; CT; CLC; MARO	04_02_24_11	367	(19)
	62.5 uM; HOMD; CT; CLC; MARO	04_02_24_12	69	(19)
	62.5 uM; HOMD; CT; CLC; MARO	04_03_17_14	56	(19)
Bleomycin	1.0 or 4.0 ug/mL; HAP; CT; PGRO	-	231	(20)
	0.01 U/mL; HOMD; AT; CLC; MARO	Bleo	50	(21)
	1.7 ug/mL; HOMD; CT; CLC; MARO	04_05_18_01	343	(19)
	1.7 ug/mL; HOMD; CT; CLC; MARO	04_05_18_02	403	(19)
	1.13 ug/mL; HOMD; CT; CLC; MARO	04_05_18_03	428	(19)
	1.13 ug/mL; HOMD; CT; CLC; MARO	04_05_18_04	499	(19)
Camptothecan	100 ug/mL; HETD; CT; CLC; MARO	Camptothecan	62	(22)
	30 uM; HAP; CT; CLC; MARO	Camptothecan	200	(23)
	30 ug/mL; HOMD; CT; CLC; MARO	04_03_25_01	35	(18, 19)
	30 ug/mL; HOMD; CT; CLC; MARO	04_03_30_01	39	(18, 19)
	30 ug/mL; HOMD; CT; CLC; MARO	04_05_12_10	87	(18, 19)
	30 ug/mL; HOMD; CT; CLC; MARO	04_07_16_08	216	(18, 19)
	250 uM; HOMD; AT; CLC; MARO	CPTa	16	(21)
	5 ug/mL; HOMD; CT; CLC; MARO	CPTc	10	(21)
Carboplatin	250 uM; HOMD; CT; CLC; MARO	03_08_27_11	53	(19)
	500 uM; HOMD; CT; CLC; MARO	03_09_11_03	36	(19)
	700 uM; HOMD; CT; CLC; MARO	04_02_24_06	271	(19)
	15 mM; HOMD; CT; CLC; MARO	04_03_09_01	77	(18, 19)
	15 mM; HOMD; CT; CLC; MARO	04_03_25_02	174	(19)
	15 mM; HOMD; CT; CLC; MARO	04_03_30_02	56	(18, 19)
CdtB	HOMD; CT; PGRO	-	61	(24)
Cisplatin	66 uM; HOMD; CT; CLC; MARO	summary_hom_cisplatin_66	102	(25)
	125 uM; HOMD; CT; CLC; MARO	summary_hom_cisplatin_125	23	(25)
	250 uM; HOMD; CT; CLC; MARO	summary_hom_cisplatin_250	49	(25)
	500 uM; HOMD; CT; CLC; MARO	summary_hom_cisplatin_66	138	(25)
	31.25 uM; HETD; CT; CLC; MARO	summary_cisplatin_31.25	4	(25)
	62.5 uM; HETD; CT; CLC; MARO	summary_cisplatin_62.5	0	(25)
	125 uM; HETD; CT; CLC; MARO	summary_cisplatin_125	6	(25)
	1.0 mM; HOMD; AT; CLC; MARO	Cis1	79	(21, 26)
	0.2 mM; HOMD; AT; CLC; MARO	Cis4	43	(21, 26)
	600 uM; HETD; CT; CLC; MARO	Cisplatin	29	(22)
	170 uM; HAP; CT; CLC; MARO	Cisplatin	71	(23)
	125 uM; HOMD; CT; CLC; MARO	01_09_18_04	27	(19)
	133.2 uM; HOMD; CT; CLC; MARO	03_02_05_06	82	(19)
	133.2 uM; HOMD; CT; CLC; MARO	03_02_05_07	147	(19)
	125 uM; HOMD; CT; CLC; MARO	03_03_20_09	421	(19)
	125 uM; HOMD; CT; CLC; MARO	03_03_20_10	389	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_03_20_11	289	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_03_20_12	319	(19)

	31.25 uM; HOMD; CT; CLC; MARO	03_03_26_11	459	(19)
	31.25 uM; HOMD; CT; CLC; MARO	03_03_26_12	376	(19)
	500 uM; HOMD; CT; CLC; MARO	03_04_04_03	342	(18, 19)
	500 uM; HOMD; CT; CLC; MARO	03_04_04_04	399	(18, 19)
	250 uM; HOMD; CT; CLC; MARO	03_04_04_05	216	(19)
	250 uM; HOMD; CT; CLC; MARO	03_04_04_06	236	(19)
	125 uM; HOMD; CT; CLC; MARO	03_04_04_07	154	(19)
	125 uM; HOMD; CT; CLC; MARO	03_04_04_08	152	(19)
	66 uM; HOMD; CT; CLC; MARO	03_04_04_11	292	(19)
	66 uM; HOMD; CT; CLC; MARO	03_04_04_12	272	(19)
	125 uM; HOMD; CT; CLC; MARO	03_08_14_14	105	(19)
	125 uM; HOMD; CT; CLC; MARO	03_10_20_03	70	(19)
	500 uM; HOMD; CT; CLC; MARO	04_02_24_05	110	(18, 19)
	500 uM; HOMD; CT; CLC; MARO	04_03_17_06	117	(18, 19)
DMAEC	240 mM; HOMD; CT; CLC; MARO	04_03_09_06	177	(18, 19)
	240 mM; HOMD; CT; CLC; MARO	04_03_17_15	262	(18, 19)
Etopside	1mM; HAP; CT; PGRO	-	11	(27)
Hydroxyurea	50-150mM; HAP; CT; PGRO	-	288	(28)
	50 mM; HOMD; CT; CLC; MARO	03_11_06_09	334	(19)
	25 mM; HOMD; CT; CLC; MARO	03_11_06_10	247	(19)
	50 mM; HOMD; CT; CLC; MARO	03_11_21_04	201	(19)
	100 mM; HOMD; CT; CLC; MARO	03_11_21_05	537	(19)
	50 mM; HOMD; CT; CLC; MARO	04_02_24_04	106	(19)
	100 mM; HOMD; CT; CLC; MARO	04_03_09_03	301	(19)
	200 mM; HOMD; CT; CLC; MARO	04_03_09_04	298	(19)
	20 mM; HETD; CT; CLC; MARO	Hydroxyurea	31	(22)
	100 mM; HOMD; CT; CLC; MARO	HU	39	(21)
	20 mM; HAP; CT; CLC; MARO	Hydroxyurea	101	(23)
Ionizing Radiation	80 krad; HOMD; AT; PGRO	-	135	(29)
	200 Gy Cs137; HOMD; AT; CLC; MARO	IR	49	(21, 30)
Mechlorethamine	20 uM; HOMD; AT; CLC; MARO	Mech	27	(21)
	62.5 uM; HOMD; CT; CLC; MARO	02_12_18_13	418	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_02_05_04	193	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_02_05_05	198	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_03_20_07	367	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_03_20_08	288	(18, 19)
	31.25 uM; HOMD; CT; CLC; MARO	03_08_27_12	175	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_08_27_13	32	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_09_11_05	34	(19)
	31.25 uM; HOMD; CT; CLC; MARO	03_10_20_12	174	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_09_16	173	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_10_05	153	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_10_07	119	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_10_09	97	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_19_01	174	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_19_02	147	(19)
	62.5 uM; HOMD; CT; CLC; MARO	03_12_19_04	204	(19)

	62.5 uM; HOMD; CT; CLC; MARO	03_12_19_06	124	(19)
	62.5 uM; HOMD; CT; CLC; MARO	04_01_21_09	113	(18, 19)
	62.5 uM; HOMD; CT; CLC; MARO	04_01_21_11	125	(19)
	62.5 uM; HOMD; CT; CLC; MARO	04_01_21_13	164	(19)
	62.5 uM; HOMD; CT; CLC; MARO	04_01_21_15	103	(19)
Melphalan	800 uM; HOMD; AT; CLC; MARO	Mel	12	(21)
	250 uM; HOMD; CT; CLC; MARO	03_09_10_01	38	(19)
	500 uM; HOMD; CT; CLC; MARO	03_09_11_02	34	(19)
	2000 uM; HOMD; CT; CLC; MARO	03_11_20_11	390	(19)
Mitomycin c	1 mM; HOMD; CT; CLC; MARO	03_01_28_05	517	(19)
	1 mM; HOMD; CT; CLC; MARO	03_02_19_03	870	(19)
	1 mM; HOMD; CT; CLC; MARO	03_02_19_04	452	(19)
	1 mM; HOMD; CT; CLC; MARO	03_11_21_03	1068	(19)
	1 mM; HOMD; CT; CLC; MARO	04_02_24_02	277	(18, 19)
	1 mM; HOMD; CT; CLC; MARO	04_03_17_03	345	(19)
	1 mM; HOMD; CT; CLC; MARO	04_03_17_08	946	(19)
	1 mM; HOMD; CT; CLC; MARO	04_03_25_03	166	(19)
	1 mM; HOMD; CT; CLC; MARO	04_03_30_03	154	(18, 19)
	0.5 mM; AT; CLC; MARO	MMC	16	(21, 26)
	0.15 uM; HAP; CT; CLC; MARO	Mitomycin C	378	(23)
MMS	0.035%; HAP; CT; PGRO	-	100	(31)
	0.001%/0.01%; HOMD; CT; CLC; MARO	-	144	(32)
	0.01-0.03% MMS; HAP; CT; PGRO	-	1403	(33)
	0.002%; HETD; CT; CLC; MARO	MMS	130	(22)
	0.03%; HOMD; CT; CLC; MARO	MMS	24	(21)
	0.004%; HAP; CT; CLC; MARO	MMS	246	(23)
	0.002%; HOMD; CT; CLC; MARO	04_01_14_08	490	(19)
	0.002%; HOMD; CT; CLC; MARO	04_02_24_03	47	(18, 19)
	0.004%; HOMD; CT; CLC; MARO	04_03_09_02	183	(19)
	0.002%; HOMD; CT; CLC; MARO	04_03_17_04	122	(19)
	0.002%; HOMD; CT; CLC; MARO	04_03_17_09	337	(19)
	0.002%; HOMD; CT; CLC; MARO	04_03_25_04	48	(18, 19)
4NQO	0.2-0.5 ug/mL; HAP; CT; PGRO	-	786	(33)
	0.0313 uM; HOMD; CT; CLC; MARO	04_08_03_07	145	(18, 19)
	0.0313 uM; HOMD; CT; CLC; MARO	04_08_05_01	111	(18, 19)
Oxaliplatin	4 mM; HOMD; CT; CLC; MARO	04_02_24_07	220	(18, 19)
	4 mM; HOMD; CT; CLC; MARO	04_03_17_07	197	(18, 19)
	1 mM; HOMD; CT; CLC; MARO	03_09_10_02	66	(19)
	4 mM; HOMD; CT; CLC; MARO	03_11_20_10	435	(19)
	10 mM; AT; CLC; MARO	Oxa	146	(21, 26)
Psoralen	62.5 uM; HOMD; CT; CLC; MARO	03_08_14_11	200	(19)
	0.5 uM irradiated; HOMD; CT; CLC; MARO	04_02_24_13	391	(19)
	0.5 uM; HOMD; CT; CLC; MARO	04_02_24_14	76	(19)
	0.5 uM irradiated; HOMD; CT; CLC; MARO	04_03_17_10	90	(18, 19)
	0.5 uM irradiated; HOMD; CT; CLC; MARO	04_03_17_11	138	(19)
	0.5 uM; HOMD; CT; CLC; MARO	04_03_17_13	31	(19)
	0.5 uM irradiated; HOMD; CT; CLC; MARO	04_03_25_05	55	(18, 19)

Streptozotocin	2 mM; HOMD; CT; CLC; MARO	03_10_08_13	226	(18, 19)
	2mM; HOMD; CT; CLC; MARO	04_03_17_16	85	(18, 19)
	2mM; HOMD; CT; CLC; MARO	04_03_25_07	64	(18, 19)
	2mM; HOMD; CT; CLC; MARO	04_07_16_09	204	(18, 19)
t-BuOOH	0.50-1.25 mM; HAP; CT; PGRO	-	439	(33)
Teniposide	500 uM; HOMD; CT; CLC; MARO	02_12_18_12	52	(19)
Tirapazamine	250/300 uM; HOMD; AT; CLC; MARO	TPZ	44	(21)
Topotecan	20 uM; HOMD; AT; CLC; MARO	Tpt	3	(21)
Ultraviolet Light	UVB 3400 J/m2; HOMD; AT; CLC; MARO	UVB	19	(21, 34)
	UVC 200 J/m2; HOMD; AT; CLC; MARO	UVC	26	(21, 34)
	UVA 36/288 J/cm2; HOMD; AT; CLC; MARO	UVA	13	(21)
	110-270 J/cm2; HOMD; AT; CLC; MARO	-	160	(32)
	40-125 J/m2; HAP; AT; PGRO	-	284	(33)

*HAP = haploid strains, HOMD = homozygous diploid strains, HETD = heterozygous diploid strains, AT = acute treatment; CT = chronic treatment; CLC = competitive liquid culture, MARO = microarray readout, PGRO = plate growth readout

†Mutations from the new and reanalyzed data in Brown *et al*. (21) were included if their $\log_2$ ratio were < -0.69. Mutations from the new and reanalyzed data in Hillenmeyer *et al*. (19) and the data in Lum *et al*. (22) and Parsons *et al.* (23) were included if their p scores were < 0.01. Other studies included those mutations directly reported as causing sensitivity. Numbers of genes reflect merging of dubious genes with any verified gene that they overlap.

Table S4. Computational test for recovery of known genes removed from the original dataset.

Screen	Reported Genes Recovered	Reported Genes Missed	Ref.
CAN1 mutator GCR Screen	7/8 (*TSA1*, *SKN7*, *YAP1*, *SHU1*, *SHU2*, *ELG1*, and *YDL162C*)	1/8 (*RNH203*)	(2)
pif1 GCR Screen	8/11 (C*SM2*, *ELG1*, *MMS4*, *RAD5*, *RAD18*, *TSA1*, *CDC50*, and *YDL162C*)	3/11 (*ALO1*, *ESC1*, and *UFO1*)	(35)
Alternative GCR Screen	13/16 (*RAD27*, *MRE11*, *SGS1*, *RAD6*, *SLX8*, *SLX5*, *WSS1*, *ESC2*, *RMI1*, *RML2*, *RAD5*, *TOP3*, and *THP2*)	3/16 (*BUD16*, *ZIP1*, and *PDX3*)	(36)

Table S5. Mutations that did not cause significant increases in GCR rates.

Genotype*	Systematic Name	RDKY Number	Cluster	Number of DNA Damaging Screens	Rate*
Wild type	-	3615	n.a.	n.a.	3.5×10^{-10} (1)
ard1::HIS3	*yhr013c*	6210	4	12	Low ($<1.4\times10^{-9}$)
ccr4::HIS3	*yal021c*	6228	4	41	Low ($<1.4\times10^{-9}$)
cla4::G418	*ynl298w*	6274	2	46	Low ($<6.3\times10^{-10}$)
csm1::HIS3	*ycr086w*	6230	2	23	Low ($<8.5\times10^{-10}$)
ctf19::HIS3	*ypl018w*	6425	29	6	8.3×10^{-10} (2.4)
ctk1::HPH	*ykl139w*	6412	32	7	Low ($<8.2\times10^{-10}$)
ctk2::HIS3	*yjl006c*	7603	32	9	Low ($<2.0\times10^{-9}$)
dbf2::HIS3	*ygr092w*	7005	11	7	Low ($<1.2\times10^{-9}$)
doa1::HIS3	*ykl213c*	6240	3	23	Low ($<1.1\times10^{-9}$)
dpb3::TRP1	*ybr278w*	5034	21	5	Low ($<3.6\times10^{-10}$)
dpb4::G418	*ydr121w*	7022	21	12	4.7×10^{-10} (1.3)
eaf1::HIS3	*ydr359c*	6483	4	22	Low ($<1.7\times10^{-10}$)
get1::HIS3	*ygl020c*	6218	1	23	Low ($<1.0\times10^{-9}$)
get2::HIS3	*yer083c*	6220	1	48	Low ($<9.8\times10^{-10}$)
hpr1::7615	*ydr138w*	7615	32	11	Low ($<2.7\times10^{-10}$)
hst4::HIS3	*ydr191w*	7529	50	9	2.1×10^{-10} (0.6)
hur1::HIS3	*ygl168w*	6204	1	15	Low ($<6.7\times10^{-10}$)
lge1::HIS3	*ypl055c*	6393	3	24	Low ($<8.2\times10^{-10}$)
lte1::HIS3	*yal024c*	6135	2	39	Low ($<1.1\times10^{-9}$)
mlh1::G418	*ymr167w*	6653	58	2	Low ($<6.9\times10^{-10}$)
mms22::HIS3	*ylr320w*	6200	4	16	Low ($<2.1\times10^{-9}$)
mon2::HIS3	*ynl297c*	7606	32	12	Low ($<6.3\times10^{-9}$)
mrc1::TRP1	*ylc061c*	5105	4	37	Low ($<6.0\times10^{-10}$)
msn5::HIS3	*ydr335w*	7067	3	26	Low ($<8.0\times10^{-10}$)
nas6::HIS3	*ygr232w*	6232	69	2	Low ($<1.1\times10^{-9}$)
nat1::HIS3	*yld040c*	6208	4	4	Low ($<1.8\times10^{-9}$)
npl3::HIS3	*ydr432w*	7604	32	12	Low ($<5.1\times10^{-10}$)
npt1::HIS3	*yor209c*	6141	32	9	Low ($<8.7\times10^{-10}$)
nup120::HIS3	*ykl057c*	6640	7	7	Low ($<1.6\times10^{-9}$)
pap2::TRP1	*yol115w*	5652	12	7	8.5×10^{-10} (2.4)
pby1::HIS3	*ybr094w*	6224	32	4	Low ($<1.0\times10^{-9}$)
pop2::G418	*ynr052c*	7478	4	14	Low ($<1.1\times10^{-9}$)
pph3::HIS3	*ydr075w*	6238	16	69	Low ($<9.2\times10^{-10}$)
pso2::G418	*ymr137c*	7480	72	70	7.2×10^{-10} (2.1)
psy2::G418	*ynl201c*	7027	60	55	Low ($<5.4\times10^{-10}$)
rad2::HIS3	*ygr258c*	7060	74	75	Low ($<4.5\times10^{-10}$)
rad23::HIS3	*yel037c*	6133	33	71	Low ($<6.4\times10^{-10}$)
rad61::HIS3	*ydr014w*	6222	29	43	Low ($<9.0\times10^{-10}$)
rnr4::G418	*ygr180c*	7485	32	14	Low ($<7.8\times10^{-10}$)
rpn4::HIS3	*ydl020c*	6214	3	6	Low ($<7.4\times10^{-10}$)
rts1::HIS3	*yor014w*	7004	8	9	Low ($<7.6\times10^{-10}$)
sap30::HIS3	*ymr263w*	7028	3	20	Low ($<8.2\times10^{-10}$)
sgo1::HIS3	*yor073w*	7009	32	5	Low ($<1.0\times10^{-9}$)
slx4::HIS3	*ylr135w*	7522	55	63	5.0×10^{-10} (1.4)
spt4::G418	*ygr063c*	6651	32	9	Low ($<5.9\times10^{-10}$)
swi6::G418	*ylr182w*	7488	4	21	Low ($<1.4\times10^{-9}$)
tpp1::G418	*ymr156c*	7490	53	11	6.3×10^{-10} (1.8)
ubc4::HIS3	*ybr082c*	6139	5	2	Low ($<1.0\times10^{-9}$)
ubp6::HIS3	*yfr010w*	6236	3	16	Low ($<8.2\times10^{-10}$)
ump1::HIS3	*ybr173c*	7010	32	5	Low ($<8.0\times10^{-10}$)
vac7::HIS3	*ynl054w*	7007	‡	2	Low ($<1.9\times10^{-9}$)
wss1::HIS3	*yhr134w*	6137	53	27	Low ($<9.4\times10^{-10}$)
ypt6::HIS3	*ylr262c*	7029	1	12	Low ($<6.4\times10^{-10}$)

*Deletions were constructed in the RDKY3615 [*MATa leu2Δ1 his3Δ200 trp1Δ63 ura3-52 ade2Δ1 ade8 lys2ΔBgl hom3-10 hxt13::URA3*] background.

† "Low" indicates rates that were below the detection limits of the performed measurements and were not pursued further as they were not substantially higher than the wild-type rate. Parentheses indicate fold increase relative to the wild-type rate.

‡ Gene falling into the unclustered group

Table S6. Random simulations of DNA damaging agent screens give remarkably consistent probabilities regardless if all 155 are simulated or if the screens are divided into two groups, probability calculated from independent simulations, and combined.

n1	*n2*	p_{in}	p_{out}	$p_{combined} = p_{in} \times p_{out}$
0 of 22	4 of 133	0.646811	0.097414	0.063009
1 of 22	3 of 133	0.284740	0.189376	0.053923
2 of 22	2 of 133	0.059708	0.273161	0.016310
3 of 22	1 of 133	0.007935	0.259851	0.002062
4 of 22	0 of 133	0.000750	0.122287	0.000092
		Summed from above		**0.135396**
		Probability for 4 hits in 155 screens		**0.1354**